DOCUMENT RESUME

ED 260 077 SP 026 503

AUTHOR Smith, Charles W.
TITLE Verbal Behavior and Classroom Practice.
PUB DATE 22 Aug 84
NOTE 36p.; Paper presented at the International Conference on Thinking (Cambridge, MA, August 19-23, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classroom Techniques; Cognitive Processes; *Critical Thinking; Learning Strategies; *Questioning Techniques; *Teacher Behavior; Teacher Made Tests; *Test Use; *Verbal Communication

ABSTRACT

The literature on the verbal behavior of teachers and their methods of testing students was reviewed. Studies included those on structuring (focusing attention on the topic), soliciting (including questions, commands, imperatives, and requests), responding (answering questions), and reacting (clarifying, synthesizing, expanding, or making positive or negative comments). In analyzing these studies, focus was upon the use of Bloom's taxonomy for structuring educational objectives, developing oral questioning strategies, and writing test items. Bloom's cognitive domain included, in ascending order: knowledge, comprehension, application, analysis, synthesis, and evaluation, each representing a progressive step from lower-order to higher-order thinking. The conclusion drawn from examining study results was that teacher oral questioning, teacher-made test items, instructional materials, and those tests available commercially have, for the most part, a heavy emphasis on rote recall and recognition. Most test items failed to assess analysis, synthesis, or evaluation. It is proposed that teacher education programs needs to give emphasis to questioning strategies in their student teaching seminars and in methods courses. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED260077

Verbal Behavior and Classroom Practice

Paper presented at

The International Conference on Thinking

at

The Graduate School of Education

Harvard University

Cambridge, Massachusetts

August 22, 1984

by

Dr. Charles W. Smith

Associate Professor

Northern Illinois University

DeKalb, Illinois 60115

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C.W. Smith

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SP 026503

Most verbal activity in the classroom, according to Bellack, Kliebard, Hyman, and Smith (1966) is teacher-talk. Bellack et al. (1966) analyzed taped sessions of 15 high-school classes from New York City that were studying a unit on international trade. The categories used included structuring (focusing attention on the topic), soliciting (including questions, commands, imperatives, and requests), responding (answering questions), and reacting (clarifying, synthesizing, expanding, or making positive or negative comments).

Some of the findings included:

1. Teachers were responsible for 70-75% of the verbal activity in the classroom.
2. Stating facts and explaining principles accounted for 50-60% of the total discourse. Lesser amounts of time were devoted to defining terms or to expressing and justifying personal opinions.
3. Most teachers' moves were those of soliciting (about 47% and reacting (about 39%). Teachers did comparatively little structuring (about 8%) or responding (about 6%).
4. Most students' moves were those of responding (about 65%). Students asked few questions (about 11%), resulting in the small percentage of teacher response noted above.
5. Though the discussion topic included trade policies, very little use was made of personal opinions about trade policies by either teachers or pupils.

6. Teacher questions averaged one and one-half lines of typed script. Student responses tended to be shorter.
7. Students performed the task expected by the solicitor in 92% of the responding moves.
8. Teacher questioning was looked upon as requiring a single student response. Student questioning was looked upon as expecting a teacher response. Seldom was pupil-pupil interaction anticipated.

Table 1

Pedagogical Move	Percentage of Teachers' Moves	Percentage of Pupils' Moves
Soliciting	46.6	11.3
Reacting	39.2	15.1
Structuring	7.7	1.8
Responding	5.5	65.4
Not Codable	1.0	6.4
	100.0	100.0

Amidon and Hough (1967) state that nearly 70% of classroom instructional time is utilized as either teacher-talk or pupil-talk. The fact that teacher questioning is a primary teaching technique is undisputed. Gall (1970) stated that "the value of focusing on teachers' questions is that they are the basic unit underlying most methods of classroom teaching" (p. 719). Hyman (1979), with regard to teacher questioning, asserts "No other single pedagogical technique yields so much return for your effort" (p.5). Aschner (1961) says the teacher might be called a "professional question maker" and asserts

that it is by asking questions and analyzing their responses that teachers measure and evaluate the progress of their students. What we must not assume is that "teacher talk" necessarily constitutes the most effective learning atmosphere. Even if such practice imparts a great abundance of "facts," information does not necessarily equate with "learning." In the social studies handbook developed by the Ventura County Superintendent of Schools, it is stated that knowledge per se, past or present, is insufficient in today's changing world. Students need to develop thinking competencies that will aid them in solving current and future problems.

Over the years many types of observational systems have been used to record and analyze teacher verbal behavior. Gall (1970) described eight systems that have been used for classifying verbal behavior. The conclusion reached was that "Bloom's Taxonomy (see Tables 2 and 3) best represents the commonalities that exist among systems" (p. 710). Over the years it has become the framework used to measure the variety of objectives and levels of questioning used by teachers and has been used as a guide to help upgrade questioning strategies. The cognitive domain of Bloom's Taxonomy (1956) was developed by college examiners who conceived of the taxonomical framework as early as 1948. It was designed as a sequential and cumulative hierarchy with levels of thinking from simple to complex; from concrete to abstract. Each of the higher levels is thought to be dependent upon knowledge, skills, and abilities of lower levels. This presentation will stress the use

of Bloom's taxonomy for structuring educational objectives, developing oral questioning strategies, and writing test items. Few innovations have had the impact upon education as has the taxonomy in its use by researchers, psychometricians, curriculum specialists, and teachers. Blumberg et al. (1932) have stated: "Current psychological theories of learning and memory support the use of a hierarchical structure in classifying examination questions. . . . When higher taxonomic level questions are answered correctly it can be assumed that the student has mastered the prerequisite basic forms of learning as well as the intellectual skills required to answer the question" (p. 2).

What levels of cognitive thought are encouraged by the kinds of questions that teachers ask? Gall (1970) analyzed a number of studies that had been conducted during the 50-year period from 1920-1970. The following was stated:

> It is reasonable to conclude that in a half century, there has been no essential change in the types of questions which teachers emphasize in the classroom. About 60% of teachers' questions require students to recall facts; about 20% require students to think; and the remaining 20% are procedural. (p. 713)

More recently, Trachtenberg (1974) analyzed the questions on nine sets of world history test materials using Bloom's Taxonomy. Over 61,000 items were analyzed. The percentage of questions assessing at the knowledge level was approximately 63%. There were about 35% of

the items at the comprehension level. More than 98% of the items were found to be at the two lowest levels of knowledge and comprehension. Less than 1% per category was included at the application, analysis, synthesis, and evaluation levels.

In 25 schools in Lebanon, Billek (1974) taped lessons in biology, chemistry, and physics. Teachers also submitted examinations to correspond with the lessons. Class lectures and exam items were analyzed to determine the relationship between the subject matter taught and the level of test items used for assessment. The nature of the subject being taught (biology, chemistry, physics) appeared to have no relationship to the level of test items used. The questioning at grade seven and grade ten did not differ significantly with respect to cognitive level. Test item analysis revealed approximately 72% knowledge items, 21% comprehension factors and 7% components at the application level. No items assessed analysis, synthesis, or evaluation. It was concluded that science exams in Lebanon did not assess critical thinking.

When West African examinations in biology, chemistry, and physics were analyzed, Black (1960) found similar results. Observations were made at two different times during the school year. Biology classes averaged approximately 83 and 94% knowledge items and approximately 17 and 6% comprehension factors. No other levels were assessed. In different chemistry classes, knowledge items approximated 64 and 66%; comprehension items approximated 15 and 26%; and application items

approximated 21 and 8% of questioning. No levels higher than application were assessed. In various physics classes, knowledge items averaged in the 48 and 56% range; comprehension in the 28 and 18% range; and application in the 24-26% range. Again no questions were utilized at the analysis, synthesis, or evaluation levels.

A study of teacher-made tests in ninth-grade courses by Davis and Pfeiffer (1965) revealed that at least half of the questions in college preparatory, business, and prevocational courses were rote oriented. Tests of civics and biology included 100% knowledge items. In the general math, algebra, general science, French I, advanced French, home economics and business training courses, there were no test questions asked that assessed analysis, synthesis, or evaluation. The world history class included synthesis level questions as well as the three lower levels. English classes were the only classes observed that consistently included questions at the analysis, synthesis, and evaluation levels.

Insert Figure 1 about here

Gallagher and Aschner (1963) used categories based on the operations of the intellect, as described by Guilford (cognitive memory, convergent thinking, divergent thinking, evaluative thinking, and routine) to analyze teacher-pupil verbal behavior with gifted children. Tape recordings were made of five consecutive junior high school classes in social studies, mathematics, science, and English.

Their analysis of the social studies data revealed that the teacher used 52% cognitive memory, 20% convergent thinking, 6% divergent thinking and 22% evaluative thinking. Both teachers and students used more questions and statements in the cognitive memory and evaluative response categories and utilized less levels of convergent and divergent thinking. Slight increases in teacher use of divergent questioning resulted in large increases in the number of student divergent responses. Gallagher and Aschner (1963) state that the kind of thinking engaged in by pupils is determined by the kinds of questions asked by teachers. Age would not seem to be a major determining factor. Klein (1972), based upon research findings, indicated general acceptance for the idea that written tests could be developed for 7-9 year olds that could measure separately and distinctly the various levels of Bloom's Taxonomy.

Fast (1974) found less usage of knowledge items than most. In analyzing 12 tests with over 950 items, the investigator found 40% were categorized at the knowledge level, 25% at the comprehension level, 25% at the application level, and 10% at the analysis level. No synthesis questions were asked, and evaluation items were virtually absent.

Hare and Pulliam (1980) studied the questioning behavior of elementary teachers using Guszak's categories of literal questions (recognition, recall, and translation), inferential questions (conjecture and explanation), and evaluation questions. Guszak (1967),

in citing previous research, stated ". . . teachers appear to equate reading-thinking skills with the most narrow of literal comprehension skills" (p. 227). The results of the Hare and Pulliam study were in very close agreement with Guszak's (1967) original data. Literal questions were most frequently used: 71% in Guszak's study and approximately 74% in the Hare and Pulliam study. Guszak found about 14% of questioning to be at the inferential level; Hare and Pulliam noted approximately 11% in this category. Evaluation questions comprised about 15% of the questions in both studies. It was noted however, that the evaluation responses in their research were mostly of the "yes" or "no" variety.

Konen et al. (1981) indicated that questioning at lower cognitive levels was associated with lower student achievement and also had the effect of discouraging students from thinking at higher levels. Farnsworth and Daines (1978) concluded ". . . that the higher the cognitive level reached and maintained by the teacher during an instructional session the higher the mean performance of pupils following an instructional session" (p. 73) Interestingly, Mills and Rice (1980) and Riegle (1976), besides suggesting teacher-training in questioning, also suggest student familiarization with question classification systems. It was thought that this approach would help to insure more questioning/responding at the higher levels of cognition.

Are standardized and commercial tests any better than teacher-

made tests in terms of taxonomical level assessed? Pancella (1971) analyzed 41 tenth-grade biology tests that had been prepared by commercial publishers, test bureaus, and the Biological Curriculum Study. Of the 2,689 items analyzed, approximately 72% assessed at the knowledge level, 15% through comprehension, 11% by application, 1% by analysis, and a fractional percent at the synthesis and evaluation levels. In several tests, 100% of the items were at the knowledge level. Very few of the tests had more than 10% of the items above the comprehension level.

Insert Figure 2 and 3 about here

The message that comes through loud and clear is that teacher oral questioning, teacher-made test items, instructional materials, and those tests available commercially have, for the most part, a heavy emphasis on rote recall and recognition. Some more limited attempts are made at assessing comprehension and application. Most test items failed to assess analysis, synthesis, or evaluation.

Teachers need to be good consumers when purchasing standardized tests and other teaching aids. Tests must be analyzed to see that they meet local objectives. In addition, it is prudent to review the items to examine the taxonomical level being assessed.

A well trained and highly qualified teacher may proclaim critical thinking, reflective thinking, creativity, problem-solving, and making judgments as learning outcomes. When such is the case, that

teacher's classroom discourse, tests, and instructional materials should reflect such emphasis. Not even all specially trained educators apply what they should know about multi-dimensional kinds of thinking. It certainly cannot be assumed that those who have produced company exams, standardized tests, or commercially prepared instructional materials have done so with a knowledge of an application of Bloom's six levels of cognition — or of any system other than informational recall. Synthesis and evaluation items are generally absent from multiple-choice examinations (Fast, 1984; Cox, 1965; Stoker & Kropp, 1964). It is the responsibility of the informed instructor to insure the inclusion in teacher-made tests and written objective exams of questions requiring at least the first four levels of thinking in Bloom's Taxonomy. It is also the instructor's responsibility to find additional ways to assess thinking in the dimensions of synthesis and evaluation.

Newton (1978) stresses the importance of the assimilation and processing of information and of evaluating ideas--critical reading and thinking. Reading teachers stress literal comprehension without always developing the thought processes of analysis, synthesis, and evaluation that are necessary for critical reading and thinking.

Teacher-made tests reflect the "real objectives" (my emphasis) held by teachers (Pfeiffer & Davis, 1965). In classrooms of highly qualified educators, teachers state their objectives and make them known to students prior to teaching a unit so as to provide direction

to the learning process. The degree to which teacher objectives provide a good match to the test items, a matter of content validity, varies from teacher to teacher.

A test blueprint or table of specifications is a useful tool in the test development process. This two-way table, with content areas on one axis and educational objectives on the other axis, is a great aid in insuring content validity. By taking into account the percentage of emphasis for each major objective, and the amount of time spent on each content area, the teacher can judge the number of items per objective/content area that will best match the instructional emphasis to the testing emphasis. The taxonomy can be a useful tool in specifying one's objectives on the test blueprint.

Reliability coefficient is not a term reserved for psychometricians. Yet many educators would not be able to respond if asked, "What was the reliability coefficient of your last major exam?" Marshall and Hales (1971) indicate that the reliability of most teacher-made tests is probably in the area of +.20 +.40. Many tests, they state, have reliabilities approaching zero. Traditional norm-referenced tests are dependent upon spreading scores for obtaining adequate reliability coefficients (Smith, 1973). It is necessary, therefore, to write questions utilizing a wide range of difficulty levels. Because of its hierarchical arrangement, the taxonomical structure can provide guidance in this situation. Knowledge items for the most part will be the easiest items on the test. Comprehension

items will be more difficult. Application of what one knows and comprehends will test at yet a higher level.

Fast (1974), Cox (1965), Stoker and Kropp (1964) found that as the taxonomical level increased, there was a concurrent increase in the difficulty level of the items. There was an inverse relationship between taxonomical level and measurement score.

It is appropriate for teachers to analyze their unit tests. On the keyed copy of an examination a note can be made of the taxonomical level of each question, thereby noting over-emphasis or neglect at any level. The total composition of the test in terms of taxonomical levels will readily reveal itself and the instructor's purpose. Of course, each test need not measure at all taxonomical levels. The instructor may wish to assess some higher levels of the taxonomy through means other than the written objective test. A multiple-choice format, for instance, could put great constraints on measuring synthesis and evaluation. If essay exams are used for this purpose, care must be taken that the items are actually requiring students to perform at specified higher levels. Many essay examinations, as presently written, measure little more than rote knowledge and comprehension.

Students soon recognize the kind of test question used in a given class and adjust their thinking accordingly. Their future study habits are heavily influenced by the kinds of items to which they have been exposed.

A wider distribution of scores than would have occurred otherwise is often obtained when a deliberate attempt is made to teach and assess at the various cognitive levels. This is probably true because many students have become accustomed to tests that assess primarily at the rote level of learning. Questions at a higher level are more likely to have a "best answer" than a "right answer" choice. Such items may be more argumentative but, if well written, will assess at a plane higher than recall of rote knowledge. The wider spread of scores obtained on norm-referenced tests are, from a statistical standpoint, to be desired. Reliability is enhanced by obtaining greater differences between adjacent scores in a distribution. Such differences are more likely to be reflective of knowledge gain when the scores are spread out. To the extent that scores become restricted, differences between adjacent scores may more likely reflect chance differences in achievement. Keep in mind that the greater spread of scores does not have to insure failures in the group. How we treat the spread of scores, gradewise, is up to the individual teacher.

The testing situation can become more humanized when teachers attempt to facilitate student success on examinations (Smith, 1973). One way to do this is by providing students with many non-graded formative exercises in preparation for the graded summative exercises (Smith, 1982). These kinds of exercises give students concrete examples of teacher expectations.

Regarding discrimination, Cox (1965) had found knowledge items to be least discriminating and application items to be most discriminating. Past (1974) had found comprehension items most discriminating and analysis items least discriminating. Though the data here is conflicting, it does indicate that whatever the discrimination index is for a set of items, it is probably wise to select the most discriminating items within taxonomical categories rather than selecting the most discriminating items from the total item pool.

Some thought might well be given to the rate of questioning in the classroom. In a study (Hoetker, 1968) of nine junior high school English classes verbal behavior of teachers was analyzed over a five-day period, covering 45 hours of instruction. The mean rate of questioning was 1.68 questions per minute (q/m). When seat-work time was omitted, and teacher-pupil talk time was recalculated, questioning was 3.22 q/m. When routine matter time was omitted, and questioning was confined only to subject-matter related talk, the recalculated figure became 5.17 q/m or a question every 11.8 seconds. In a ninth-grade low ability class, questioning was 10.72 q/m or a question every 5.6 seconds. With questioning at this rate, one has to wonder about the response time allocated for answering. Adults take 14 seconds to respond to an analogy question on a standardized test (Locking, 1975). With teacher questioning occurring at the frequency indicated above, one also has to be concerned about the level of questioning. When

teacher questions are occurring every 6 or 12 seconds, what level of response is required? Is it likely to be anything other than rote? When questions are occurring at this pace, is it likely that students are asked to analyze, to come up with unique solutions, or to make judgments and offer support for them?

Lapse time or wait-time is a further consideration. How soon after initially asking a question should one repeat the question, rephrase it, ask a different question, or call on another student?

Atwood and Stevens (1976), in working with junior and senior high school science teachers, found average lapse time to be one second. About 75% of the items were knowledge oriented, 25% comprehension related, and 5% application associated. Arnold et al. (1973) worked with science and social studies student teachers in primary and intermediate grades and found response time to be about two seconds, with 61% of the questions being memory level. Rowe (1974), in a six-year study analyzed 300 tape recordings of instruction and found average wait-time to be one second. In addition, there was only one second of lapse time after student responses before teachers reacted to the response, asked another question, or initiated another topic or question. Rowe (1974) conducted a study wherein training was utilized to help increase initial wait-time to 3-5 seconds. Over 900 tapes of those who had such training were analyzed. The results included the following:

1. response length increased from 7 words to 28 words;

2. the number of unsolicited but appropriate responses increased from 3 to 37;
3. failure to respond decreased;
4. confidence, as evidenced by a decrease in inflected responses, increased;
5. speculative responses increased from 2 to 11;
6. pupil-pupil interaction increased;
7. evidence-inference statements increased;
8. the number of student questions increased;
9. the number of responses from lower achievers increased; and
10. students used a wider variety of responses (structuring, soliciting, reacting).

With training, teachers decreased the number of questions asked from 7-10 per minute to only 2-3 per minute. The kind of questioning changed also. Rhetorical questions were utilized about the same; informational questions decreased; and leading and probing questions increased. It appeared that teacher expectations of lower achievers changed as the extended wait-time permitted more of the slower achievers the opportunity to respond.

Those who are in teacher education programs need to give emphasis to questioning strategies in their student teaching seminars and in methods courses. The same questioning techniques need to be emphasized with in-service teachers. Pre-service and in-service teachers need a realization of the impact of varying types

of questions and need practice and reinforcement in implementing their use in the classroom.

Forty-four secondary social studies student teachers used the Teacher-Pupil Question Inventory to analyze questioning behaviors in junior and senior high classes in a study by Davis and Tinsley (1967). Questions were analyzed in nine categories (memory, translation, interpretation, application, analyses, synthesis, evaluation, affectivity, and procedure). Both student-teachers and pupils asked more memory questions than all other categories combined. Second highest usage was questions at the translation and interpretation levels. No other cognitive levels appear to have been operational. The levels of questions asked by student-teachers and responses by pupils correlated +.90. The data regarding levels of questioning being used were in general agreement with that of Gallagher (1965) and Davis and Pfeiffer (1965).

Willson (1973) used an interaction analysis system in Grades 1-8 to provide self-evaluation of questioning and responding. The investigator found self-evaluation helpful in improving teachers' cognitive level, as reflected through their questioning. Teacher questioning, in turn, increased students' cognitive level.

Riley (1978) trained elementary student teachers in the use of a classification hierarchy called the Question Category System for Science. Its categories included cognitive memory, convergent thinking, divergent thinking, evaluative thinking, and non-

thinking. The mean cognitive level of questions used by student teachers trained in questioning behavior was significantly higher than for the control group that received no such training.

Anderson (1972) reported on teacher-made questions used in research articles in the Journal of Educational Psychology and the American Educational Research Journal for the period from January 1964 to February 1971. Anderson claimed that two-thirds of the investigators gave no rationale for their choice of questions asked on tests they developed. Most of the others were thought to have given this concern only superficial treatment.

We teachers talk a lot, and we ask a lot of questions. If teaching is talking and questioning, if it is to be expert teaching, it must include talk and questioning about more than rote facts. It must include talk and questions at the higher levels of cognition if students are to think and to learn at those levels. As a qualified, dedicated teacher, I must ask myself a few questions about my techniques.

1. Have I identified a range of significant behavioral goals for my students?
2. Are students aware of my goals of instruction?
3. Do materials that I use support and provide for the accomplishment of the same levels of objectives as do the goals I espouse?
4. Is my choice of test items representative of the content

area?

5. Do I conscientiously choose an array of oral discussion questions to encourage varying levels of thinking?
6. Do I have a rationale for my choice of questions?
7. Do I develop test items that will best measure each of my objectives?
8. Is my pacing of questioning consistent with levels of cognition that I am encouraging?
9. Do I know the reliability of my major examinations?
10. Do I utilize the higher levels of the cognitive domain and encourage my students to do the same?

All of the above are important in our teaching and assessment. In addition teachers need to be sure that their tests and workshoeets are done with great care regarding correct spelling, capitalization, punctuation, and syntax. Content accuracy needs to be assured. The test must be constructed utilizing accepted principles of test construction. A parent sometimes forms an impression of a teacher by the only tangible evidence that has been seen--tests and worksheets that students have taken home. We want to look good. Our tests must look good.

Probably the particular type of questioning categorization system used does not matter so much. What is important is that teachers become acutely aware of the verbal strategies that they are using, that they understand the kinds of thought processes encouraged, and

that they continually strive to facilitate student growth toward independence by using a variety of questioning levels. In addition, teaching effective questioning to our students will be helpful to them in assuming responsibility for their own learning.

I conclude with a statement by Suchman (cited in Sitkoff, Whitehouse, & Hodgson) which stated, "Only when teachers and pupils understand and use better techniques of questioning as part of the learning process will each pupil's real potential be achieved." Again in conclusion is a statement of my own. Effective questioning and evaluation measures with a high degree of validity are costly in terms of preparation and utilization time, but lack of good measures is more costly in terms of loss of direction toward stated objectives.

References

Adams, G. S. (1964). Measurement and evaluation in education, psychology, and guidance. New York: Holt, Rinehart and Winston.

Amidon, E. J., & Hough, J. B. (Eds.) (1967). Interaction analysis: Theory, research and application. Reading, MA: Addison-Wesley.

Aschner, M. J. (1961). Asking questions to trigger thinking. NEA Journal, 50(6), 44-46.

Atwood, R. K., & Stevens, J. T. (1976). Relationships among question level, response level and lapse time: Secondary science. School Science and Mathematics, 76(3), 249-254.

Beatty, R. (1975). Reading comprehension skills and Bloom's taxonomy. Reading World, 15(2), 101-108.

Bellack, A. A., Kliebard, H. M., Hyman, R. T., & Smith, F. L., Jr. (1966). The language of the classroom. New York: Teachers College Press.

Billeh, V. Y. (1974). An analysis of teacher-made science test items in light of the taxonomic objectives of education. Science Education, 58(3), 313-319.

Bloom, B. S. (Ed.) (1956). Taxonomy of educational objectives: The Classification of educational goals. Handbook I. Cognitive domain. New York: Longman, Green and Company.

Blumberg, P., Alshuler, M. D., & Rezmovic, V. (1982). Should taxonomic levels be considered in developing examinations? Educational and Psychological Measurement, 42(1), 1-7.

Bradley, R. C. (1966). Structuring questions--A teacher's major teaching tool: The art of questioning. Arizona Teacher, 54, 14, 29.

Carin, A. A., & Sund, R. B. (1971). Developing questioning techniques. Columbus, OH: Charles E. Merrill.

Cox, R. C. (1965). Item selection techniques and evaluation of instructional objectives. Journal of Educational Measurement, 2(2), 181-185.

Davis, O. L., & Tinsley, D. C. (1967). Cognitive objectives revealed by classroom questions asked by social studies student teachers. Peabody Journal of Education, 45(1), 21-26.

Farnsworth, B. J., & Daines, D. (1978). A qualitative study of cognitive instructional levels as related to pupil achievement. Reading Improvement, 15(1), 69-73.

Fast, K. V. (1974). An analysis and classification of the ACS-NSTA high school chemistry achievement tests using Bloom's taxonomy-cognitive domain. Science Education, 58(1), 17-21.

Gaier, E. L. (1952). Selected personality variables and the learning process. Psychological Monographs, 66(17), 1-28.

Gall, M. D. (1970). The use of questions in teaching. Review of Educational Research, 40(5), 707-721.

Gallagher, J. J. (1965). Expressive thought by gifted children in the classroom. Elementary English, 42(5), 559-568.

Guszak, F. J. (1967). Teacher questioning and reading. The Reading Teacher, 21(3), 227-234.

Hare, V. C., & PUlliam, C. A. (1980). Teacher questioning: A verification and an extension. Journal of Reading Behavior, 12(1), 69-72.

Hoetker, J. (1968). Teaching questioning behavior in nine junior high school English classes. Research in the Teaching of English, 2(2), 99-106.

Hunkins, F. P. (1968). The influence of analysis and evaluation questions on achievement in sixth-grade social studies. Educational Leadership, 25(4), 326-332.

Hunkins, F. P. (1972). Questioning strategies and techniques. Boston: Allyn and Bacon.

Hyman, R. T. (1979). Strategic questioning. Englewood Cliffs, NJ: Prentice-Hall.

Klein, M. F. (1972). Use of taxonomy of educational objectives (cognitive domain) in structuring tests for primary school pupils. Journal of Experimental Education, 40(3), 36-50.

Koren, S., Cohen, R., & Solman, R. (1981). A levels-of-processing analysis of Bloom's taxonomy. Journal of Educational Psychology, 73(2), 202-211.

Marshall, J. C., & Hales, L. W. (1971). Classroom test construction. Readings, MA: Addison-Wesley.

Mills, S. R., & RIce, C. T. (1980). The correspondence between teacher questions and student answers in classroom discourse. The Journal of Experimental Education, 48(3), 194-204.

Pancella, J. R. (1971). Cognitive levels of test items in commercial biology examinations. Rockville, MD: Montgomery County Public Schools. (ERIC Document Reproduction Service No. ED 052 045)

Pfeiffer, I., & Davis, O. L., Jr. (1965). Teacher-made examinations--What kinds of thinking do they demand? National Association of Secondary School Principals Bulletin, 49(302), 1-22.

Riegle, R. P. (1976). Classifying classroom questions. Journal of Teacher Education, 27(2), 156-161.

Riley, J. P. (1978). Effects of studying a question classification system on the cognitive level of preservice teachers' question. Science Education, 62(3), 333-338.

Riley, J. P. (1980). A comparison of three methods of improving preservice science teachers' questioning knowledge and attitudes toward questioning. Journal of Research in Science Teaching, 17(5), 419-424.

Rowe, M. B. (1974). Wait-time and rewards as instructional variables, their influence on language, logic and fate control: Part one-Wait time. Journal of Research in Science Teaching, 11(2), 81-94.

Sanders, N. M. (1966). Classroom questions: What kinds? New York: Harper & Row.

Sitkoff, S., Whitehouse, L. H., & Hodgson, F. (1967). The art of questioning in science: Summary and implications. Los Angeles City Schools, Division of Instructional Planning and Review.

Smith, C. W. (1973). Criterion-referenced assessment in contrast to norm-referenced measurement. Marquette University Education Review, 4(2), 23-32.

Smith, C. W. (1974). Questioning behavior that facilitates cognitive development in reading. DeKalb, IL: Northern Illinois University. (ERIC Document Reproduction Service No. ED 098 528)

Smith, C. W. (1980). A self-evaluation of teacher-made tests. College Student Journal Monograph, 14(4), 1-12. (Available from the author for $5.25).

Smith, C. W. (1982). The role of formative and summative evaluation in assessing student achievement. Eastern Kentucky University Educational Review, 5(2), 48-53.

Smith, C. W. (1982). Evaluating your classroom testing. Illinois Guidance and Personnel Association Quarterly, 84, 24-31.

Smith, C. W. & Cluff, R. B. (19). Humanizing the testing aspect of evaluation. DeKalb, IL: Northern Illinois University. (ERIC Document Reproduction Service No. ED 076 498)

Stoker, H. W., & Kropp, R. P. (1964). Measurement of cognitive processes. Journal of Educational Measurement, 1(1), 39-42.

Trachtenberg, D. (1974). Student tasks in text material: What cognitive skills do they tap? Peabody Journal of Education, 52(1), 54-57.

Willson, I. A. (1973). Changes in mean levels of thinking in grades 1-8 through use of an interaction analysis system based on Bloom's taxonomy. Journal of Educational Research, 66(9), 423-429.

Table 1

Pedagogical Move	Percentage of Teachers' Moves	Percentage of Pupils' Moves
Soliciting	46.6	11.3
Reacting	39.2	15.1
Structuring	7.7	1.8
Responding	5.5	65.4
Not Codable	1.0	6.4
Total	100.0	100.0

Table 2

Bloom's Cognitive Domain*

1.00 KNOWLEDGE

- 1.10 Knowledge of specifics
  - 1.11 Knowledge of terminology
  - 1.12 Knowledge of specific facts
- 1.20 Knowledge of ways and means of dealing with specifics
  - 1.21 Knowledge of conventions
  - 1.22 Knowledge of trends and sequences
  - 1.23 Knowledge of classifications and categories
  - 1.24 Knowledge of criteria
  - 1.25 Knowledge of methodology
- 1.30 Knowledge of the universals and abstractions in a field
  - 1.31 Knowledge of principles and generalizations
  - 1.32 Knowledge of theories and structures

2.00 COMPREHENSION

- 2.10 Translation
- 2.20 Interpretation
- 2.30 Extrapolation

3.00 APPLICATION

4.00 ANALYSIS

- 4.10 Analysis of elements
- 4.20 Analysis of relationships
- 4.30 Analysis of organizational principles

5.00 SYNTHESIS

- 5.10 Production of a unique communication
- 5.20 Production of a plan, or proposed set of operations
- 5.30 Derivation of a set of abstract relations

6.00 EVALUATION

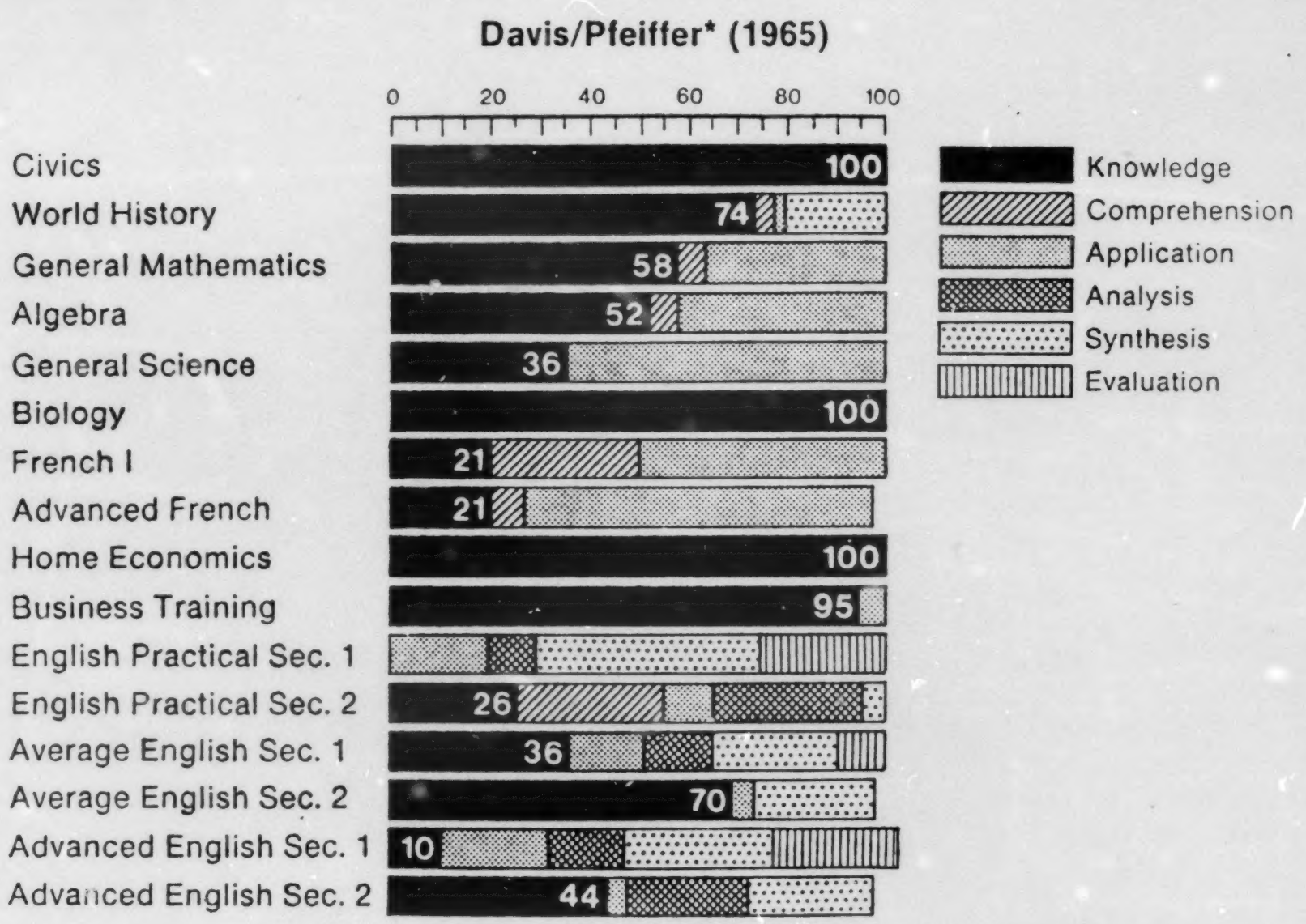
- 6.10 Judgments in terms of internal evidence
- 6.20 Judgments in terms of external criteria

*Adapted from "Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I. Cognitive Domain" by B. S. Bloom (Ed.) New York: Longman, Green and Co., 1956.

Table 3

Major Levels of Bloom's Taxonomy

6.00	Evaluation...is the highest level of the taxonomy. It includes making a value judgment about a product, idea, or solution using specified criteria. A judgment might be made regarding feasibility, accuracy, effectiveness, artistic quality, etc.	Highest Level
5.00	Synthesis...requires putting elements together to form a unique whole. Creativity is emphasized at this level.	
4.00	Analysis...involves breaking down a novel communication so that it may be more easily understood. It may involve identification of the parts, a study of the relationships among the parts, or an analysis of organizational principles.	
3.00	Application...involves the use of ideas, rules, procedures, or methods in a novel problem-solving situation. The learner selects those understandings and processes thought to be necessary for problem solution. Rules, formulas, laws, and theories may be applied.	
2.00	Comprehension...involves an understanding of the literal message conveyed through a map, chart, graph, cartoon, poem, story, etc. This may involve translation (paraphrasing or changing from one language form to another), interpretation (explanation or summarization), or extrapolation (extending trends beyond the data provided).	
1.00	Knowledge...is defined as the recall or recognition of previously learned material (facts, terminology, categories, methods, generalizations, theories). Knowledge may be merely rote recall or the material may have been only slightly modified from the way in which it was learned. Understanding of the material is not essential at this lowest level of the taxonomy.	Lowest Level



*Numbers as abstracted from the original manuscript

Figure 1.

Reported Levels of Teacher Questioning

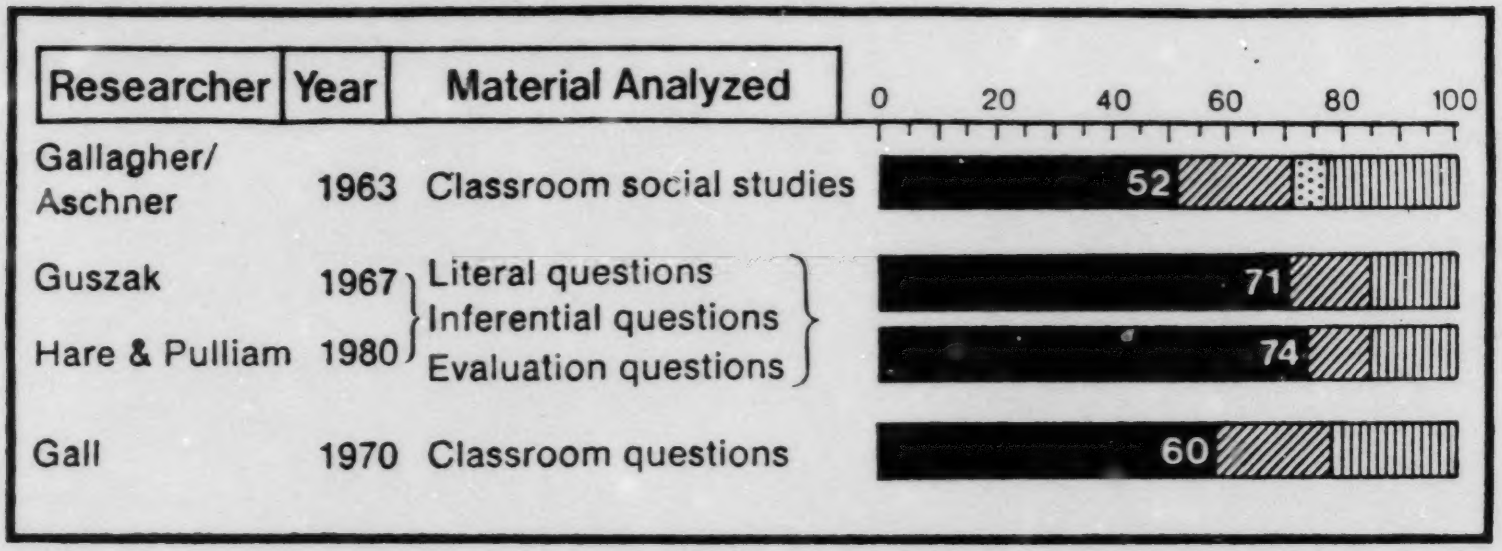


Figure 2.

Reported Levels of Teacher Questioning

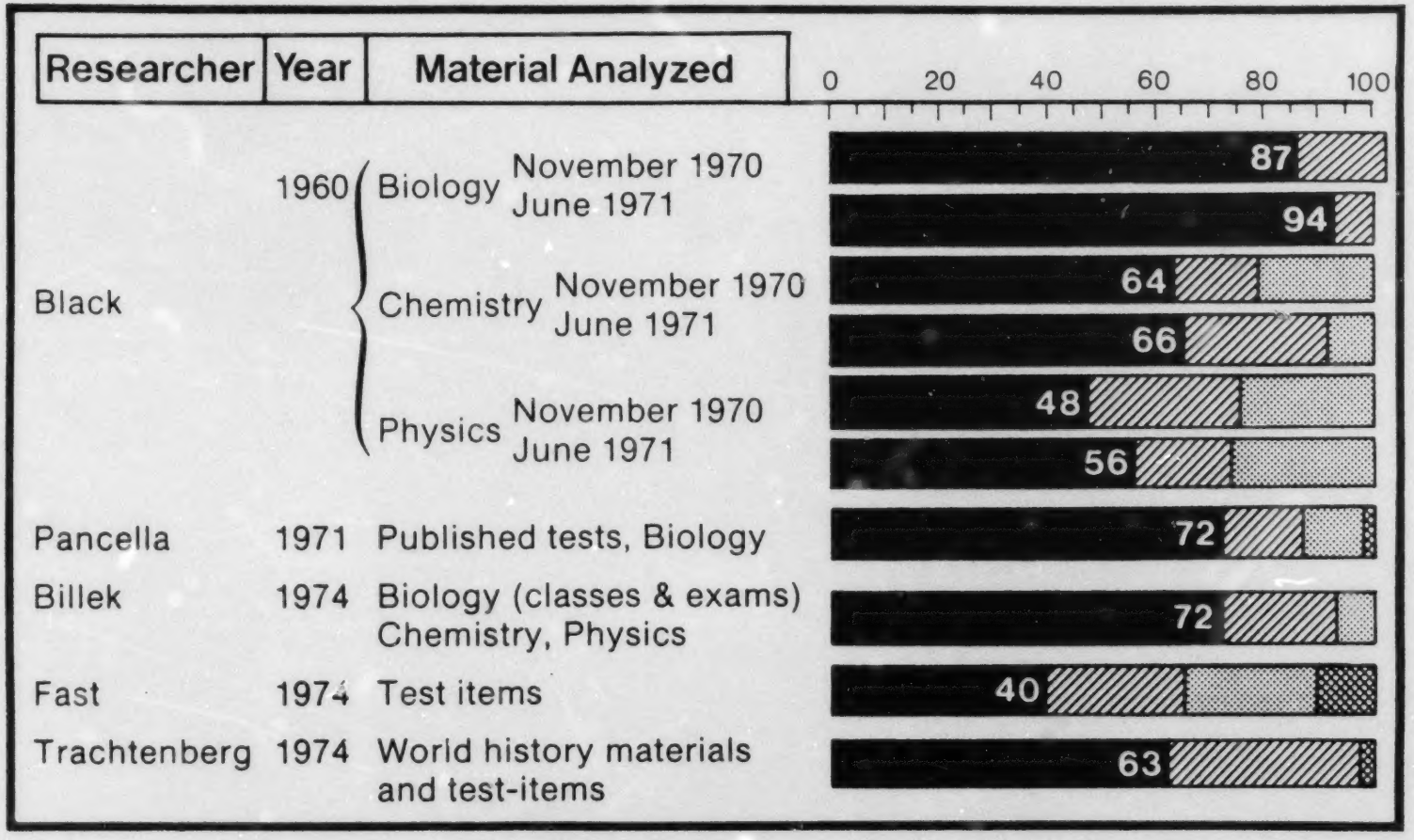


Figure 3.